Simulating Human Visual Perception in Nighttime Illumination

Ning Zhou, Weiming Dong, Wang Jiaxin, Jean-Claude Paul

TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 0?/?? pp?-??
Volume 14, Number 1, February 2009

Simulating Human Visual Perception in Nighttime Illumination

ZHOU Ning (周 宁)[1], DONG Weiming (董未名)[2], WANG Jiaxin (王家廞)[1,**], Paul Jean-Claude[1,2]

1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;
2. INRIA Lorraine, Nancy 54506, France

Abstract: This paper presents an image-based algorithm for simulating the visual adaptation of the human visual system to various illuminations, especially in dark nighttime conditions. The human visual system exhibits different characteristics depending on the illumination intensity, with photopic vision in bright conditions, scotopic vision in dark conditions, and mesopic vision between these two. A computational model is designed to simulate multiple features of mesopic vision and scotopic vision, including the chromaticity change, luminance change, and visual acuity loss. The system uses a source image under bright illumination as input. Then assuming that the viewer has already adapted to the new conditions, the color spectrum of the input image is reconstructed to replace the source with modifications of the chromaticity and the luminance of the relighted scene. A bilateral filter is used to simulate the visual acuity loss. The model parameters have clear physical meanings and can be obtained from experimental data to achieve realistic results.The algorithm can be used not only for visual perception simulation, but also as a day-for-night tool to produce realistic nighttime images from daytime images.

Key words: visual perception; scotopic vision; mesopic vision; visual acuity; bilateral filter

Introduction

The illumination at nighttime is quite complicated. It usually includes both skylight and artificial lighting such as car headlights and streetlamps. Visual perception experiments have shown that the human visual characteristics in such conditions differ from the characteristics in bright daylight conditions. At nighttime, the human visual perception tends to have lower saturation, a blue-shift of the intensity, and some visual acuity loss. These phenomena are caused by the different responses of the retina, which automatically adapts to different illuminations.

The film industry has used a technique called Day-for-Night[1] for a long time to produce nighttime effects during the daytime with blue filters and under-exposure. Similar post-processing methods are used in computer graphics to simulate nighttime visual perception[2]. General image toning techniques such as Durand and Dorsey[3] and Lischinski et al.[4] can also accomplish visually-plausible Day-for-Night transformation. However, as image tone mapping tools are not specifically designed for nighttime perception simulation, to give visually-plausible effects, they usually require time-consuming manual parameter tuning.

Though the biological research on the vision adaptation mechanism has not reached a final conclusion, empirical models have been set up on the basis of measurement data. Pattanaik et al.[5] presented a fast visual adaptation operator useful for interactive real-time applications. Shin et al.[6] proposed a applicable model for mesopic vision based on experimental data. Khan and Pattanaik[7] simulated the blue shift effect based on the rod-cone interaction theory. Haro et

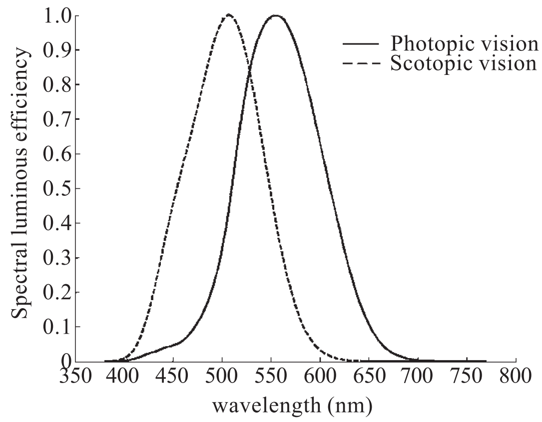
Received: 2007-09-11
** To whom correspondence should be addressed.
E-mail: wjx@mail.tsinghua.edu.cn; Tel: 86-10-62796826

al.[8] presented a comprehensive digital Day-for-Night algorithm taking into account multiple features for realistic results. Because of their photorealistic results, many of them are used in device display and rendering result correction[9]. There have also been algorithms integrated with renderer to perform night atmosphere rendering[10]. The disadvantage of these physically-based methods is that they usually contain a complicated parameter system, which are not convenient for users to make further image appearance manipulation.

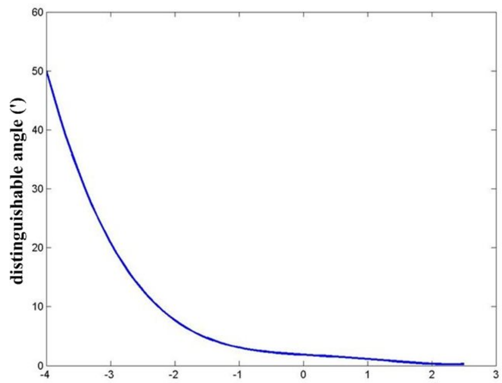
This paper presents a computational algorithm, which uses spectrum reconstruction as the bridge between measurement data in the wavelength domain and the input image in the tristimulus RGB space to take the advantage of ample measurement data presented in previous works for parameter configuration. The algorithm is integrated in an image-based system to simulate the human visual perception in nighttime illumination, in particular, very dim illuminations, in which the human visual system functions as mesopic vision. The system provides users with extensive control of the results by allowing them to select illuminations of arbitrary intensity and spectrum power distributions. It then automatically calculates the human visual perception for the scene from the source image. The system workflow is divided into three steps. First, the system reconstructs the retinal response spectra from the source image. Then the responses of cones and rods to the selected illumination are estimated based on measurement data. At last, a bilateral filter is applied to simulate the detail loss. The algorithm is not limited to nighttime conditions, but also can be used for other illuminations.

1 Human visual system

The human visual system is able to perceive light over a broad range of intensities, from 10^{-4} to 10^{6} cd/m^2. Its main light perception apparatus retina has two different types of photoreceptor cells: cones and rods. The cones are responsible for color vision in bright light, while the rods play an important role in dark environments. In a bright environment, the cones perceive colors and fine details, while in a dark environment the rods perceive the light when the cones do not function. The rods are much more sensitive to light than the cones, but they have lower visual acuity and no color perception[11]. Thus, visual perception with rods only is colorless. The response of the cones is called photopic vision, while that of the rods is scotopic vision. The standard CIE spectral luminous efficiency curves[12] of them are shown in Fig. 1a. Photopic vision has a sensitivity peak at around 555 nm (yellow-green), while scotopic vision has a sensitivity peak at around 510 nm (blue-green). The sensitivity peak moves towards shorter wavelength (blue) direction as the illumination intensity decreases. This is referred as the Purkinje phenomenon[11].



(a) Spectral luminous efficiency curves



(b) Visual acuity curve

Fig. 1 Characteristics of cones and rods[12]

In practical nighttime environments, artificial lightings cause the retinal response to be neither typical photopic vision nor typical scotopic vision, but the combined mesopic vision where cones and rods cooperate.

Besides the color response, the illumination also affects the spatial resolution, temporal resolution, and lightness resolution of the human visual system. The spatial resolution is called visual acuity, usually defined as the reciprocal of the visual angle of the minimum distinguishable detail. The visual acuity decreases with the illumination intensity as shown in Fig. 1b. The lightness resolution is represented by the local contrast where the ratio of the minimum distinguishable lightness difference ΔL to the background lightness L approaches a constant at low intensities. According to Weber's law[12], for intensities of 0.1~1.0×10^3 cd/m^2, $\frac{\Delta L}{L}=0.02\sim0.03$; when L decreases to 10^{-4} cd/m^2, $\frac{\Delta L}{L}$ increases to 0.5; theoretically, in very dark conditions, when $L\to0$, $\frac{\Delta L}{L}\to\infty$. The temporal resolution will not be discussed here since this paper only considers the final stable adaptation result.

2 Algorithm

The system workflow is shown in Fig. 2. The system starts with a source image taken in a bright environment as the initial input. The target illumination intensity is indicated by the user. The algorithm assumes that the source image is white-balance adjusted and properly exposed, so the color constancy problem does not need to be considered. The source illumination intensity is set to $E=1.0\times10^3$ by default for the following steps. Empirical models of the human visual system are used to simulate the cones and rods responses for the target illumination with a bilateral filter used to simulate the visual acuity decrease.

The main parameters in the system are:

E: source illumination intensity;

E': target illumination intensity;

σ: color matching functions;

ξ: response of the retina cells to the source illumination;

ξ': response of the retina cells to the target illumination;

α: cone adaptation function;

β: rod adaptation function;

a: weight of cone signal in the overall response;

b: weight of rod signal in the overall response.

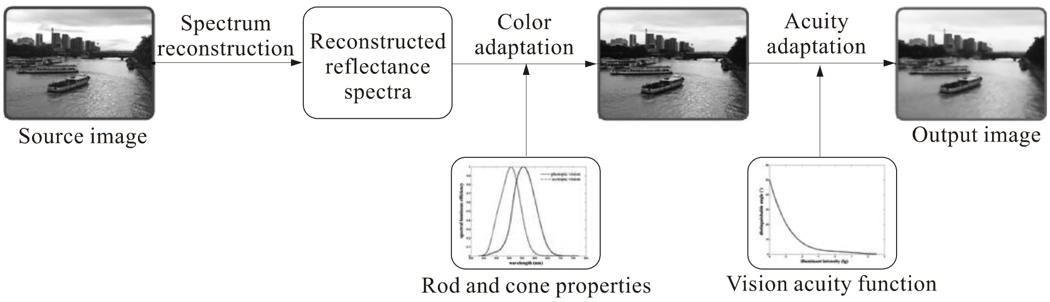


Fig. 2 System workflow

2.1 Simulating cone and rod responses

As the input image is taken in a bright environment, it is assumed to be pure cone response. Thus the input image is transformed from the RGB space to the LMS space to approximate the cone response $\xi_c=[\xi_l\ \ \xi_m\ \ \xi_s]^T$. A linear model is used to first reconstruct the color spectrum:

$$R(\lambda)=\sum_{i=\{l,m,s\}}\frac{\xi_i B_i(\lambda)}{I(\lambda)} \tag{1}$$

where B are the chosen basis functions. Assuming the source illumination type is CIE D65 by default, the B are selected to be the CIE standard color matching functions of the LMS color system[12]. Basis function design techniques can be employed to minimize the reconstruction error[13].

Then for the target illumination $I'(\lambda)$, the new cone response ξ'_c is

$$\xi'_c=\int_{\lambda_{min}}^{\lambda_{max}} R(\lambda)I'(\lambda)\boldsymbol{\sigma}'_c(\lambda)\,\mathrm{d}\lambda \tag{2}$$

where λ is the wavelength, and [λ_{min}, λ_{max}] is the range of visible light recognized by the human eyes. As the spectral efficiency curves of cones and rods are different, when the illuminant condition varies, we should correct the color matching functions σ with illuminant intensity E. We define the new color matching functions $\boldsymbol{\sigma}'_c(\lambda)$ in the target illumination as $\boldsymbol{\sigma}'_c(\lambda)=\frac{\alpha(E')}{\alpha(E)}\boldsymbol{\sigma}_c(\lambda)$ where the adaptation function α is the cone spectral efficiency curve under illuminant conditions. It can be obtained from measurement data in previous works[6].

By using $\sigma_{\mathrm{r}}(\lambda)$ instead of $\boldsymbol{\sigma}_{\mathrm{c}}(\lambda)$, the rod response ξ_{r}' in the target condition can be calculated as

$$\xi_{\mathrm{r}}' = \int_{\lambda_{\min}}^{\lambda_{\max}} R(\lambda) I'(\lambda) \sigma_{\mathrm{r}}'(\lambda) \mathrm{d}\lambda \tag{3}$$

where $\sigma_{\mathrm{r}}'(\lambda) = \dfrac{\beta(E')}{\beta(E)} \sigma_{\mathrm{r}}(\lambda)$ is the rod color matching function. The rod adaptation function β can be also obtained from measurement data[6].

2.2 Color adaptation

The final perception result is determined by the overall retina response ξ' contributed by both the cones and rods. The rod response is assumed to be an additive component to each channel in the LMS space. Thus the overall response ξ' is

$$\xi' = \begin{bmatrix} a_{\mathrm{ll}} & a_{\mathrm{ml}} & a_{\mathrm{sl}} \\ a_{\mathrm{lm}} & a_{\mathrm{mm}} & a_{\mathrm{sm}} \\ a_{\mathrm{ls}} & a_{\mathrm{ms}} & a_{\mathrm{ss}} \end{bmatrix} \begin{bmatrix} \xi_{\mathrm{l}}' \\ \xi_{\mathrm{m}}' \\ \xi_{\mathrm{s}}' \end{bmatrix} + \begin{bmatrix} b_{\mathrm{rl}} \\ b_{\mathrm{rm}} \\ b_{\mathrm{rs}} \end{bmatrix} [\xi_{\mathrm{r}}'] \tag{4}$$

where a and b represent the interactions between the cones and rods, as well as between different cone channels. Assuming no interaction between different cone channels and that the rods have a uniform influence on all the cone channels, Eq. (4) can be simplified as

$$\xi' = \begin{bmatrix} a_{\mathrm{ll}} & 0 & 0 \\ 0 & a_{\mathrm{mm}} & 0 \\ 0 & 0 & a_{\mathrm{ss}} \end{bmatrix} \begin{bmatrix} \xi_{\mathrm{l}}' \\ \xi_{\mathrm{m}}' \\ \xi_{\mathrm{s}}' \end{bmatrix} + \begin{bmatrix} b \\ b \\ b \end{bmatrix} [\xi_{\mathrm{r}}']$$

Further assuming that color match functions σ already include intensity normalization of different channels, Eq. (4) can be further simplified to

$$\xi' = a\xi_{\mathrm{c}}' + b[\xi_{\mathrm{r}}'] \tag{5}$$

where parameters a and b represent the weights of cone and rod responses in the overall perception result. They are dependent on the illumination intensity, and their values are obtained from measurement data[6].

The output image is computed with Eqs. (2), (3), and (5), and then transformed from the LMS space back to the RGB space. A sample color adaptation result in Fig. 3 shows the desaturation and Purkinje phenomenon. The result is more physically-plausible comparing to the result of Shin et al.[6] in Fig. 3c, because the algorithm is based on spectrum reconstruction rather than direct processing in a tristimulus color space. However, the spectrum reconstruction model parameters must be configured for each input image to minimize the error for the user indicated source illuminations, since illumination mismatching will affect the color adaptation accuracy.

(a) Source image (b) Our result (c) Result of Shin et al.[6]

Fig. 3 Color adaptation result. The target illumination intensity is $E' = 0.0001E$

2.3 Visual acuity adaptation

As stated in Section 1, in dim environments, both the spatial resolution and the lightness resolution of the human visual system decrease as the illumination intensity decreases, which causes detail losses in the perception result. However, the losses are not uniform blur such as the result from a Gaussian filter since the strong edges remain sharp while small details are no longer distinguishable[11].

Therefore, this system uses a bilateral filter[14] to simulate the visual acuity adaptation, which is much faster than the anisotropic diffusion method[15]. The bilateral filter is a non-linear filter with both a spatial kernel and an intensity kernel. Its result is that only pixels both spatially closed and with similar intensities have high influence. The output of a bilateral filter at pixel p is

$$I_p' = \frac{1}{k(p)} \sum_{n \in \Omega} f(n-p) g(I_n - I_p) I_n \tag{6}$$

where Ω is the neighborhood of pixel p, I is the intensity, f is the spatial filter, g is the intensity filter, and

$k(p)=\sum_{n\in\Omega} f(n-p)g(I_n-I_p)$ is the normalization term.

In practical implementation, the fast bilateral filter presented by Durand and Dorsey[16] is used, with two Gaussian functions $f(N_f,\sigma_f)$ and $g(N_g,\sigma_g)$ as the filter function, where N is the window size and σ is the attenuation speed. The size N_f of the spatial filter f represents the visual field, so it is a constant determined by the focal length of the given source image. When the focal length is not available, the image size is used to determine N_f as $N_f=\dfrac{\min(\text{width, height})}{16}$. The attenuation speed defined as $\sigma_f=\dfrac{1}{\lg E'+10}$ represents the influence of the illumination. For the intensity filter g, its size $N_g=\max_{n\in\Omega}\|I_p-I_n\|$ is the local lightness dynamic range, and $\sigma_g\propto\dfrac{\Delta L}{L}$ keeps a direct ratio with the minimum distinguishable lightness difference ΔL. According to Weber's law[12], in typical nighttime environments, $\sigma_g\approx 0.2$.

3 Results and Discussion

Some sample results are shown in Fig. 4, where the first column is the source image, and the second and the third columns are simulated visual perception result in different illumination conditions. The system performance mainly depends on the image resolution. Processing a 640×480 image costs about 12 s on a 1.8-GHz PC.

(a) Source image (b) Simulation for $E'=0.001E$ (c) Simulation for $E'=0.0001E$

Fig. 4 Sample results

One limitation of this system is that it has no special treatment for the sky, which is one of the main lighting sources during the daytime. Input scenes containing large areas of sky may have unnatural results in the output image. A preliminary sky detection step[17] can help to solve this problem by using a spatial Gaussian weight to darken the sky area. An example of this sky treatment is shown in Fig. 5. Furthermore improvements of the result quality will require specific treatments of the shadows and highlights.

(a) Source image (b) Without sky treatment (c) With sky treatment

Fig. 5 Effect of sky treatment. The target illumination intensity is $E' = 0.0001E$, which is equivalent to a nighttime environment. Without the sky treatment, the sky in (b) is unnatural bright. The result in (c) is more natural.

4 Conclusions

This paper describes an image-based method for generating the visual perception result of the human visual system for nighttime illuminations. The empirical model for the mesopic vision considers multiple visual features comprehensively. It uses a set of simple but physically-meaningful parameters, which are convenient for user control. The system is able to transform a source image taken with bright illumination into a visually-plausible result for user-indicated target illumination. The system parameters are calculated from visual perception experimental data. The algorithm can be used not only for visual perception simulation, but also as a day-for-night tool to produce realistic nighttime images from daytime images.

References

[1] Samuelson D W. Motion Picture Camera Techniques, second ed. Boston: Focal Press, 1984.

[2] Thompson W B, Shirley P, Ferwerda J A. A spatial post-processing algorithm for images of night scenes. *Journal of Graphics Tools*, 2002, **7**(1): 1-12.

[3] Durand F, Dorsey J. Interactive tone mapping. In: Eurographics Workshop on Rendering. Czech: Springer-Verlag Wien, 2000: 219-230.

[4] Lischinski D, Farbman Z, Uyttendaele M, Szeliski R. Interactive local adjustment of tonal values. In: Proceedings of SIGGRAPH 2006. USA: ACM Press, 2006: 646-653.

[5] Pattanaik S N, Tumblin J, Yee H, Greenberg D P. Time-dependent visual adaptation for fast realistic image display. In: Proceedings of SIGGRAPH 2000. USA: ACM Press, 2000: 47-54.

[6] Shin J, Matsuki N, Yaguchi H, Shioiri S. A color appearance model applicable in mesopic vision. *Optical Review*, 2004, **11**(4): 272-278.

[7] Khan S M, Pattanaik S N. Modeling blue shift in moonlit scenes using rod cone interaction. *Journal of Vision*, 2004, **4**(8): 316a.

[8] Haro G., Bertalmío M, Caselles V. Visual acuity in day for night. *International Journal of Computer Vision*, 2006, **69**(1): 109-117.

[9] Pattanaik S N, Ferwerda J A, Fairchild M D, Greenberg D P. A multiscale model of adaptation and spatial vision for realistic image display. In: Proceedings of SIGGRAPH 1998. USA: ACM Press, 1998: 287-298.

[10] Jensen H W, Premoze S, Shirley P, Thompson W B, Ferwerda J A, Stark M M. Night rendering. Technical Report UUCS-00-016. Computer Science Department, University of Utah., 2000.

[11] Hess R F, Sharpe L T, Nordby K. Night Vision. Cambridge: Cambridge University Press. 1990.

[12] Wyszecki G, Stiles W S. Color Science: Concepts and Methods, Quantitative Data and Formulae. New York: John Wiley & Sons, Inc., 1982.

[13] Sun Y. A spectrum-based framework for realistic image synthesis [Dissertation]. Simon Fraser University, Canada, 2000.

[14] Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of IEEE International Conference on Computer Vision. India: IEEE, 1998: 836-846.

[15] Barash D. A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(6): 844-847.

[16] Durand F, Dorsey J. Fast bilateral filtering for the display of high dynamic range images. In: Proceedings of SIGGRAPH 2002. USA: ACM Press, 2002: 257-266.

[17] Shotton J, Winn J, Rother C, Criminisi A. Textonboost: Joint appearance, shape and context modelling for multi-class object recognition and segmentation. In: Proceedings of the 9th European Conference on Computer Vision. Austria: Springer Berlin / Heidelberg, 2006: 1-15.